Function Estimation Employing Exponential Splines

V. Dose* and R. Fischer*

*Max-Planck-Institut für Plasmaphysik, EURATOM Association, Boltzmannstr. 2, 85748 Garching, Germany

Abstract. We introduce and discuss the use of the exponential spline family for Bayesian nonparametric function estimation. Exponential splines span the range of shapes between the limiting cases of traditional cubic spline and piecewise linear interpolation. They are therefore particularly suited for problems where both, smooth and rapid function changes occur.

Keywords: Nonparametric density estimation, splines, exponential splines

INTRODUCTION

Nonparametric function estimation is required if we wish to represent an empirical set of data (x_i, f_i) in the absence of a first principles based theory. If such a theory is available, it need not be correct and nonparametric fits may point to certain deficiencies which may trigger improvement of the theory. But even if parametric (theory based) and nonparametric fits turn out to be in close agreement, their first derivatives may differ considerably. In cases where the gradients are the information of primary importance and this applies to all transport phenomena in physics, the greater flexibility of the nonparametric approach is likely to provide the more reliable results.

EXPONENTIAL SPLINES

Consider a set of function values f_i given at support points x_i. The exponential spline function $S_i(x)$ in the interval $x_i \leq x \leq x_{i+1}$ is then given by

$$S_i(x) = \alpha_i + \beta_i(x-x_i) + \gamma_i\psi_i(x-x_i) + \delta_i\phi_i(x-x_i). \tag{1}$$

The auxiliary functions ψ_i and ϕ_i contain a stiffness parameter λ_{i+1} on the support $[x_i, x_{i+1}]$ and are given by

$$\psi_i(x) = 2\{\cosh[\lambda_{i+1}(x-x_i)] - 1\}/\lambda_{i+1}^2 \tag{2}$$

$$\phi_i(x) = 6\{\sinh[\lambda_{i+1}(x-x_i)] - \lambda_{i+1}(x-x_i)\}/\lambda_{i+1}^3 \tag{3}$$

Since the series expansions of the hyperbolic functions are

$$\cosh(\lambda z) = 1 + \frac{(\lambda z)^2}{2!} + \ldots \tag{4}$$

$$\sinh(\lambda z) = \lambda z + \frac{(\lambda z)^3}{3!} + \ldots \tag{5}$$

we realize that ψ_i and ϕ_i tend to $(x-x_i)^2$ and $(x-x_i)^3$ in the limit of λ tending to zero. In this limit $S_i(x)$ reduces to a third order polynomial on $[x_i, x_{i+1}]$ which when subject to the requirements of continuity of function, first and second derivative becomes the traditional cubic spline. It will be shown below that in the opposite limit of λ tending to infinity the nonlinear terms in x, ψ_i and ϕ_i, vanish as $1/\lambda$ and (1) reduces in this limit to a linear interpolation.

Exponential splines minimize the functional

$$\Phi(S) = \sum_i \int_{x_i}^{x_{i+1}} \left\{ |S_i''(x)|^2 + \lambda_{i+1} |S_i'(x)|^2 \right\} dx. \tag{6}$$

For vanishing λ_i this functional has a minimum if S is a cubic spline. If on the other hand the curvature term $|S_i''(x)|^2$ is negligible compared to the first derivative term the functional is minimized by a polygon which may therefore also be called a linear spline function. It is important to note that in the general case the stiffness parameters λ are different on every interval. This allows for a change of character of the exponential spline function from linear to third order polynomial on adjacent support intervals. It is this property which provides the extremely high flexibility of the exponential spline function which also offers continuity of function, first and second derivative.

The so far unknown coefficients $\alpha, \beta, \gamma, \delta$ are determined from the requirement of continuity of function, first and second derivatives at the pivotal points x_i. Continuity of function and second derivative yields already an explicit representation of the exponential spline function in terms of function values $\{f_i\}$ and second derivatives $\{M_i\}$ at the pivotal points $\{x_i\}$. Introducing the definitions

$$\begin{aligned} h_{i+1} &= x_{i+1} - x_i \\ z_i &= \lambda_{i+1}(x - x_i) \\ \mu_{i+1} &= \lambda_{i+1} h_{i+1} \end{aligned} \tag{7}$$

we obtain

$$\begin{aligned} S_i(x) = {} & \frac{x_{i+1}-x}{h_{i+1}} f_i + \frac{x-x_i}{h_{i+1}} f_{i+1} \\ & + \frac{M_i}{\lambda_{i+1}^2} \frac{\sinh(\mu_{i+1} - z_i) + (z_i/\mu_{i+1} - 1)\sinh(\mu_{i+1})}{\sinh(\mu_{i+1})} \\ & + \frac{M_{i+1}}{\lambda_{i+1}^2} \left\{ \frac{\sinh(z_i) - z_i}{\sinh(\mu_{i+1})} - \frac{z_i}{\mu_{i+1}} \frac{\sinh(\mu_{i+1}) - \mu_{i+1}}{\sinh(\mu_{i+1})} \right\} \end{aligned} \tag{8}$$

The terms involving the function values f_i and f_{i+1} represent the linear interpolation part of $S_i(x)$. The terms involving the second derivatives M_i and M_{i+1} introduce the curvature. In order to determine the so far unknown second derivatives $\{M_i\}$ in terms of the function values $\{f_i\}$ we use finally the continuity requirement for the first derivative. This yields the system of equations

$$\begin{aligned} & M_{i-1}h_i\frac{\sinh(\mu_i)-\mu_i}{\mu_i^2\sinh(\mu_i)}+ \\ & M_i\left\{h_i\frac{\mu_i\cosh\mu_i-\sinh\mu_i}{\mu_i^2\sinh\mu_i}+h_{i+1}\frac{\mu_{i+1}\cosh\mu_{i+1}-\sinh\mu_{i+1}}{\mu_{i+1}^2\sinh\mu_{i+1}}\right\} \\ & M_{i+1}h_{i+1}\frac{\sinh\mu_{i+1}-\mu_{i+1}}{\mu_{i+1}^2\sinh\mu_{i+1}} \quad = \quad \frac{f_{i+1}-f_i}{h_{i+1}}-\frac{f_i-f_{i-1}}{h_i} \end{aligned} \tag{9}$$

For N pivotal points this is a system of $N-2$ equations. The system can be closed by putting $M_1 = M_N = 0$ or by given values of the first derivative at the end points. Inspection of the coefficients of M_{i-1} and M_{i+1} shows that they vanish as $1/\lambda^2$ for $\lambda \to \infty$. The coefficient of M_i on the other hand vanishes only as $1/\lambda$ and M_i becomes proportional to λ in this limit. The terms M/λ^2 in (8) vanish therefore as $1/\lambda$ which concludes the proof that (8) converges to a polygon in the limit $\lambda_i \to \infty$ for all i since the coefficients of M_i/λ_{i+1}^2 and M_{i+1}/λ_{i+1}^2 in (8) remain smaller than two.

PERFORMANCE

The application in plasma physics which we shall exploit is the determination of profile gradients from experimentally determined profile data. In order to demonstrate the suitability of the exponential spline family for this estimation problem mock data were generated as follows. Let $g(r)$ be the gradient of the test profile

$$g(r) = -\frac{c}{1+\left|\frac{r-r_0}{\Delta}\right|^k} \tag{10}$$

where r_0 and Δ determine position and width of the function $g(r)$ and k tunes the steepness of the gradient. For small k, e.g. $k=2$, this function is bell shaped. For $k \to \infty$ it assumes the shape of a rectangle. The mock profile $G(r)$ was generated from (10) choosing $r_0 = 0.7$ and $\Delta = 0.1$ by numerical integration with posterior choice of c such that $G(0) = 1$. Twenty profile values G_i uniformly distributed between $r=0$ and $r=1$ were then converted to mock data d_i employing a Gaussian random number generator of mean G_i and shape parameter σ_i

$$\sigma_i = 0.07(0.5 + U(0,1])G_i \tag{11}$$

Profile and gradient reconstruction was afforded on a grid of 39 pivots using five support pivots for the exponential spline. Their position except for the end points $r=0$ and

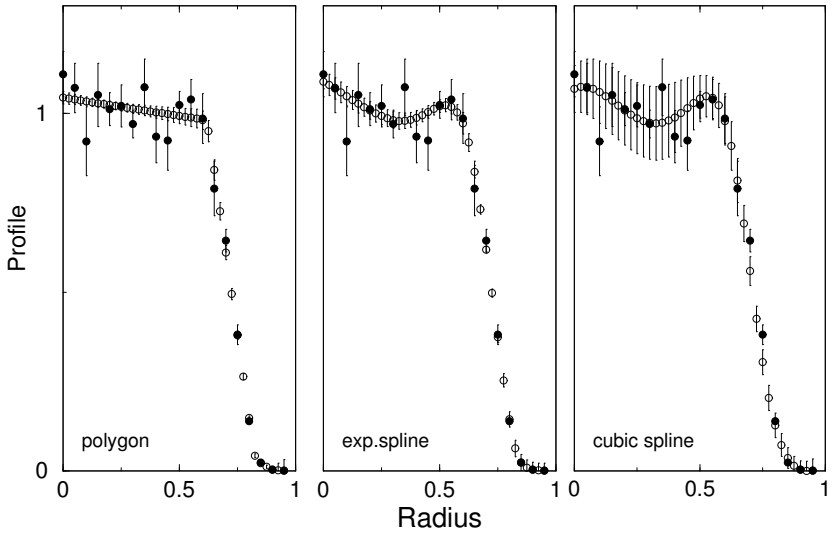

FIGURE 1. Full dots with error bars show the mock data. Open circles with uncertainties are the reconstruction. Data were generated with $k = 6$.

$r = 1$ which were kept fixed was marginalized employing simple sampling on the grid of reconstruction pivots. Marginalization was also performed for the stiffness parameters $\{\lambda_i\}$. Since they are setting the scale (7) for the variables of the problem the appropriate prior to afford this marginalization is a bounded form of Jeffreys' prior. Figure 1 shows mock data for a steepness parameter $k = 6$ as full dots and profile reconstructions with error bars as open circles for the three models polygon, exponential spline and cubic spline. The reconstruction looks quite satisfactory in all three cases apart from the region $r \leq 0.5$ in polygon reconstruction. It ignores the structure in the data introduced by the randomization and is therefore obviously too stiff for this data set. Note that both the cubic and the exponential spline account for the structure introduced by the low uncertainty points 5 to 7 and 11 to 13. More interesting is the reconstruction of the gradients shown in figure 2. The polygon reconstruction is again quite stiff and overestimates the steepness of the profile. The cubic spline on the other hand is too smooth and suffers also from a shift of the reconstructed profile (open circles) with respect to the generating profile (10). The exponential spline exhibits a nearly perfect fit for $r > 0.5$. The proper evaluation of the quality of the reconstruction is however not by visual inspection but by straightforward calculation of the prior predictive value (evidence). Table I shows the ppv in units of the natural logarithm for the polygon and the cubic spline with respect to the exponential spline. All of them are negative meaning that the exponential spline approach wins the game for all three test cases. For a smooth profile, $k = 4$, its evidence is marginally higher than the cubic spline. For a rigid profile, $k = 20$, we find that the evidence for the exponential spline is close to but slightly larger than for the polygon model. Very interesting is the intermediate case which corresponds to the reconstruction in figure 1 and figure 2. The evidence analysis singles out the exponential spline model with very high significance for this profile of intermediate stiffness.

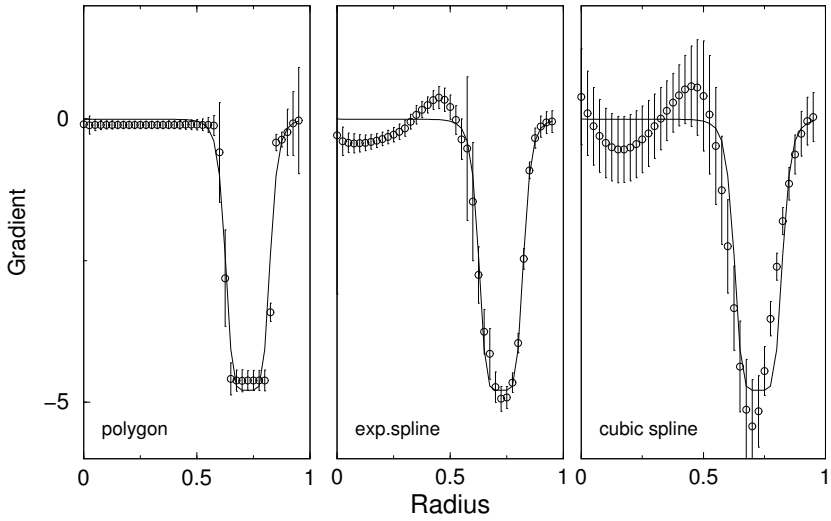

FIGURE 2. Profile gradients are shown as open circles with error bars. The continuous curve shows the data generating function (10).

TABLE 1. Prior predictive values of the cubic spline and the polygon reconstructions relative to the exponential spline in units of the natural logarithm for three different profile stiffness parameters.

k	**polygon**	**cubic spline**
4	-6.12	-0.70
6	-8.25	-10.18
20	-0.61	-14.46

CONCLUSION

The family of exponential spline functions which comprises as limiting cases polygons and cubic splines is a very versatile set for Bayesian nonparametric function estimation. The added complexity by the necessity to marginalize over the stiffness parameters $\{\lambda_i\}$ is unimportant in view of the everywhere amply available computation power.

REFERENCES

1. J. Stoer, "Einführung in die Numerische Mathematik I," 3rd edition 1979, Springer Verlag Berlin-Heidelberg-New York, p. 99.